THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 273, No. 47, Issue of November 20, pp. 31230–31236, 1998
Printed in U.S.A.

Transport and Activation of the Vacuolar Aspartic Proteinase Phytepsin in Barley (*Hordeum vulgare* L.)*

(Received for publication, June 26, 1998, and in revised form, August 26, 1998)

Stefanie Glathe‡§, Jukka Kervinen¶‖, Manfred Nimtz, Grace H. Li‡‡, Gregory J. Tobin‡‡, Terry D. Copeland¶, David A. Ashford§§, Alexander Wlodawer¶, and Júlia Costa‡§¶¶‖‖**

*From the ‡Instituto de Biologia Experimental e Tecnológica/Instituto de Tecnologia Química e Biológica, Apartado 12, 2780 Oeiras, Portugal, ¶Advanced Bioscience Laboratories-Basic Research Program and ‡‡Laboratory of Cell and Molecular Structure-Science Applications International Corp., NCI, National Institutes of Health, Frederick Cancer Research and Development Center, Frederick, Maryland 21702, **Gesellshaft für Biotechnologische Forshung, D-38124 Braunschweig, Germany, §§Glycobiology Research and Analytical Facility, Department of Biology, University of York, UK, and ¶¶Unidade de Ciências Exactas e Humanas, Universidade Algarve, 8000 Faro Portugal*

The primary translation product of barley aspartic proteinase, phytepsin (EC 3.4.23.40), consists of a signal sequence, a propart, and mature enzyme forms. Here, we describe post-translational processing and activation of phytepsin during its transport to the vacuole in roots, as detected by using metabolic labeling and immunoprecipitation. After removal of the signal sequence, the glycosylated precursor of 53 kDa (P53) was produced and further processed to polypeptides of 31 and 15 kDa (P31 + P15) and, subsequently, to polypeptides of 26 and 9 kDa (P26 + P9), 45 min and 24 h after synthesis, respectively. The processing occurred in a late-Golgi compartment or post-Golgi compartment, because brefeldin A inhibited the processing, and P53 acquired partial endoglycosidase H resistance 30 min after synthesis, whereas P15 was completely resistant. The *N*-glycosylation inhibitor tunicamycin had no effect on transport, but the absence of glycans on P53 accelerated the proteolytic processing. Phytepsin was also expressed in baculovirus-infected insect cells. The recombinant prophytepsin underwent autoproteolytic activation *in vitro* and showed enzymatic properties similar to the enzyme purified from grains. However, a comparison of the *in vitro*/*in vivo* processing sites revealed slight differences, indicating that additional proteases are needed for the completion of the maturation *in vivo*.

Aspartic proteinases (APs)[1] (EC 3.4.23) constitute one of the four superfamilies of proteolytic enzymes. They are present in a wide variety of organisms, such as viruses, fungi, plants, and animals. Common features of APs include an active site cleft that contains two catalytic aspartic acid residues (32 and 215 in pepsin), acidic pH optima for enzymatic activity, inhibition by pepstatin A, a conserved overall fold, and a preferential cleavage specificity for peptide bonds between amino acid residues with bulky hydrophobic side chains. Both intracellular and extracellular forms of APs are present in animal tissues (1, 2).

Aspartic proteinases are synthesized as inactive precursors (zymogens) in which the N-terminal propeptide is bound to the active site cleft, thus preventing undesirable protein degradation and enabling spatial and temporal regulation of proteolytic activity. Pepsinogen, the inactive precursor of stomach pepsin, needs only a drop in pH for the autocatalytic cleavage of the propeptide to result in an active enzyme (1). Procathepsin D, which is targeted to the lysosome largely via the mannose-6-phosphate receptor (3), is activated after cleavage of its N-terminal 44 amino acids, most likely by lysosomal cysteine proteinases (4). Procathepsin D is also capable of acid-dependent autoactivation *in vitro* to yield a catalytically active (pseudo)cathepsin D (5, 6). However, autocatalytic removal of the remaining 18-residue propeptide or the processing intermediate corresponding to pseudocathepsin D has not been observed *in vivo* (7).

Barley AP (*Hordeum vulgare* AP), recently renamed phytepsin (EC 3.4.23.40) (8), was originally isolated from grains in which it exists as two enzymatically active two-chain forms (9). Sequence alignment of phytepsin with animal and microbial APs shows a high degree of similarity, with the exception of an inserted domain of approximately 100 residues that is plant-specific (10–12) and very similar to that of saposins (13). The exact function of phytepsin is still controversial. Because phytepsin is an intracellular enzyme residing in leaf and root vacuoles (14) and in scutellar and aleuronal vacuole-like protein bodies in grains (15) and because it is able to cleave the C-terminal vacuolar targeting signal of barley prolectin *in vitro* (14), phytepsin may represent a cathepsin D-like enzyme from plant cells. Accordingly, phytepsin may participate in protein processing and metabolic turnover (Ref. 16 and references therein). It has recently also been observed that phytepsin may play a role in the active autolysis in plant tissues undergoing developmentally regulated programmed cell death (17).

Modification of phytepsin during its intracellular route to vacuoles involves several steps; however, the enzymology, sequence, and intracellular localization of these events is not known. The study presented here describes the mode of expression, processing, and activation of phytepsin during its transport. Furthermore, we demonstrate the autoactivation of phytepsin *in vitro* by using a recombinant enzyme expressed in insect cells.

* This work was supported by Grants PBICT/C/BIO/2025/95 and PRAXIS PCNA/P/BIO/52/96, Portugal (to J. C.), by the National Institutes of Health (NCI) Department of Health and Human Services, under contract with Advanced Bioscience Laboratories (to J. K., T. D. C., and A. W.), and by the protocol Junta Nacional de Investigacae Científica e Tecnológica/British Council 423/RU (to J. C. and D. A. A.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Recipients of fellowships BPD/4154/94 and BPD/11839/97, Portugal, respectively.

‖ Supported in part by the Academy of Finland.

‖‖ To whom correspondence should be addressed: ITQB/IBET, Apartado 12, P-2780 Oeiras, Portugal. Tel.: 351-1-4469462; Fax: 351-1-4411277; E-mail: jcosta@itqb.unl.pt.

[1] The abbreviations used are: AP, aspartic proteinase; Endo H, endoglycosidase H; PAGE, polyacrylamide gel electrophoresis; PVDF, polyvinylidene difluoride; MALDI-TOF MS, matrix-assisted laser desorption ionization time-of-flight mass spectroscopy; r-, recombinant.

EXPERIMENTAL PROCEDURES

Materials—H. vulgare (cv. Sereia) grains were purchased from the Estação Nacional de Melhoramento de Plantas, Elvas, Portugal. A purified antiserum against phytepsin was prepared as described previously (14). L-[^{35}S]methionine and L-[^{35}S]cysteine (Pro-mix™ *in vivo* cell labeling mix) and autoradiography films (Hyperfilm™-MP) were purchased from Amersham Pharmacia Biotech. Molecular weight standards for gel electrophoresis were purchased from Bio-Rad. Protein A coupled to Sepharose was purchased from Amersham Pharmacia Biotech. Endoglycosidase H (Endo H) was from Oxford GlycoSciences (Oxford, UK). Immobilon™-P polyvinylidene difluoride (PVDF) transfer membranes were purchased from Millipore Corp. (Bedford, MA). The remaining reagents were of analytical grade.

Plant Culture, Protein Extraction, and Immunoprecipitation—Barley grains were surface-sterilized with 1% (w/v) sodium hypochlorite followed by 60% (v/v) ethanol and germinated on Petri dishes containing 15 ml of 0.8% (w/v) agar for 3 days at 28 °C in the dark. For the extraction of root proteins, nondenaturing solubilization buffer (30 mM Tris, pH 7.5, 1 mM EDTA, 0.25 M sucrose, 5% (w/v) polyvinylpolypyrrolidone, 0.15% (v/v) β-mercaptoethanol) was added to the roots frozen in liquid nitrogen. Five root tips were homogenized with 0.2 ml of the buffer using a Teflon homogenizer (Sigma). The extracts were cleared by centrifugation for 5 min at 10,000 rpm, and the supernatants were used for immunoprecipitation or Western blot analysis. Protein A-Sepharose (3 mg/sample) was resuspended in immunoprecipitation buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.1% (v/v) Triton X-100, 0.05% (w/v) sodium deoxycholate, 10% (v/v) glycerol, 1% (w/v) bovine serum albumin, 1 mM EDTA) and washed twice. Rabbit antibodies were directly coupled to the Protein A-Sepharose by shaking for 20 min at room temperature. The proteins in the lysate supernatants were immunoprecipitated with aliquots of the protein A-Sepharose-antibody-protein complex for 1 h on ice. Immunoprecipitates were washed twice each with high salt (50 mM Hepes, pH 7.5, 0.5 M NaCl, 5 mM EDTA, 0.2% (v/v) Triton X-100, 0.1% (w/v) SDS), medium salt (50 mM Hepes, pH 7.5, 0.15 M NaCl, 5 mM EDTA, 0.2% (v/v) Triton X-100), and low salt (10 mM Tris, pH 7.5, 0.1% (v/v) Triton X-100) buffers (18). The complex was resuspended in 25 μl of SDS-polyacrylamide gel electrophoresis (PAGE) sample buffer, boiled for 5 min, and analyzed by SDS-PAGE (19).

Western Blot Analysis—The proteins were transferred from the gel to a PVDF membrane in a Bio-Rad semi-dry transfer cell for 1 h at 15 V. The PVDF membrane was blocked in Tris-buffered saline containing 0.05% (v/v) Tween 20 and 3% (w/v) dried lowfat milk for 45 min. The antiserum and the second antibody (alkaline phosphatase-coupled goat-anti-rabbit IgG) were used at 1/500 and 1/8000 dilutions, respectively, in blocking solution containing 1% dried lowfat milk. Detection was performed by the enhanced chemiluminescence (ECL) method (Amersham Pharmacia Biotech) or with 4-chloro-1-naphthol (20).

Metabolic Labeling and Pulse-Chase Experiments—For each time point, 5 roots from 3-day-germinated grains were cut ~1 cm from the tip, and the cut ends were dipped in 20 μl of essential B5 plant growth medium containing 4% (w/v) sucrose and 100 μCi of Pro-mix™ (L-[^{35}S]methionine/cysteine) for the indicated pulse periods. The chase was performed in 20 μl of essential B5 plant growth medium containing 4% (w/v) sucrose, 1 mM L-methionine, and 0.5 mM L-cysteine for the indicated time periods. Protein extracts were prepared from roots frozen in liquid nitrogen and were immunoprecipitated with anti-phytepsin serum. The proteins (~40 μg) were separated by SDS-PAGE in a gradient gel (10–20%). The gel was dried, and ^{35}S-labeled proteins were detected by fluorography.

Endo H Digestion of Phytepsin—Metabolic labeling and pulse-chase experiments were performed as described above, except that for each time point, 10 root tips were incubated with 200 μCi of Pro-mix™. After immunoprecipitation, the protein was dissociated from the antibodies with 40 μl of 0.1 M glycine-HCl buffer, pH 3.0, then incubated in 50 mM citrate buffer, pH 5.5, containing 1 mg/ml SDS and 0.2 M β-mercaptoethanol at 95 °C for 10 min. Phenylmethylsulfonyl fluoride, E-64, and pepstatin A were added to final concentrations of 1 mM, 10 μg/ml, and 10 μg/ml, respectively. Six milliunits of Endo H were added to a final volume of 100 μl, and the mixture was incubated at 37 °C for 18 h. The protein was then precipitated with ethanol and analyzed by SDS-PAGE in 12 or 10–20% polyacrylamide gradient gels.

Experiments with Inhibitors—Root tips (1 cm) were dipped for 1 h in 14 μl of essential B5 plant growth medium containing 4% (w/v) sucrose and inhibitors at the following concentrations: 7 μM brefeldin A_1 in methanol and 10 μg/ml tunicamycin A_1 in methanol. The control samples were incubated with the same volume of methanol added. The roots were radiolabeled as described above.

Determination of Incorporation of Radiolabeled Amino Acids into Protein—Duplicate samples of 1 μl of metabolically labeled extract were spotted on small pieces of Whatman 3MM filter paper. The papers were dried, and one of the duplicates was washed 5 times with 5 ml of 5% (v/v) trichloroacetic acid and dried. The papers were placed at the bottom of plastic vials, 3 ml of liquid scintillation mixture (Beckman) was added, and the radioactivity was counted in the carbon channel of a Beckman scintillation counter.

Expression of Phytepsin in Insect Cells—Sequences encoding the complete preprophytepsin (1863 base pairs (10)) were cloned into the *Kpn*I restriction site of the baculovirus transfer vector pBlueBac4.5 (Invitrogen Corp., Carlsbad, CA). The construct was cotransfected with *Autographa californica* multiple nuclear polyhedrosis viral DNA into *Spodoptera frugiperda* (Sf9) cells, and recombinant baculoviruses were derived using standard methodologies (21). For protein production, Sf9 suspension cultures were infected with *A. californica* multiple nuclear polyhedrosis virus-phytepsin in complete Grace's medium (about 600 ml) supplemented with 7% fetal bovine serum. Four days after infection, the conditioned culture supernatant was clarified for 20 min at 6,000 × *g* and concentrated to approximately 2% of the initial volume at 4 °C by ultrafiltration (Amicon YM30 membrane, Amicon, Danvers, MA). The concentrated supernatant was diluted 1.3-fold with cold 0.5 M sodium acetate, pH 4.0, and phytepsin was purified according to Sarkkinen *et al.* (9) by affinity chromatography on a pepstatin-agarose column, with the exception that no washing with pH 7.5 buffer was carried out and the elution was performed with 0.1 M Tris-HCl, pH 8.8, 0.2 mM dithiothreitol, 0.1 M NaCl. Further purification was performed by ion exchange chromatography on a Mono Q column (Amersham Pharmacia Biotech).

In Vitro Processing of Phytepsin—Recombinant phytepsin purified on the Mono Q column (2.2 mg/ml in 20 mM Tris-HCl, pH 8.0, 0.1 M NaCl) was mixed with an equal volume of 0.2 M incubation buffer and kept at 37 °C for up to 90 min. Incubation buffers were as follows: sodium lactate, pH 3.7, sodium acetate, pH 4.5 and 5.5, and sodium phosphate, pH 6.5. Samples were removed after various times and frozen at –70 °C. Processing products were separated by SDS-PAGE (PhastSystem, Amersham Pharmacia Biotech) and stained with Coomassie Brilliant Blue R-250 or electroblotted onto a PVDF membrane for N-terminal sequencing. The Coomassie Brilliant Blue-stained protein bands in the membrane were excised and individually sequenced on an Applied Biosystems 477A gas-phase sequencer, and the phenylthiohydantoin amino acids were identified on-line with a Model 120 analyzer (Applied Biosystems, Inc., Foster City, CA).

Proteinase Assay—Phytepsin samples were diluted with 20 mM sodium lactate, pH 3.7, and preincubated for 1 h at 37 °C, and the activity was measured according to Sarkkinen *et al.* (9) using bovine hemoglobin as the substrate at pH 3.7. The assay temperature was 37 °C. One unit of activity corresponded to the enzymatic activity that liberated trichloroacetic acid-soluble reaction products equivalent to 1 mg of bovine serum albumin in 1 h at 37 °C.

Matrix-assisted Laser Desorption Ionization Time-of-flight Mass Spectroscopy (MALDI-TOF-MS)—Glycoproteins were analyzed using a matrix of 22.4 mg of 3,5-dimethoxy-4-hydroxycinnamic acid in 400 μl of acetonitrile and 600 μl of 0.1% (v/v) trifluoroacetic acid in H_2O as the UV-absorbing material. The solubilized samples were mixed with the same volume of matrix, and 1 μl of the mixture was spotted onto the stainless steel tip and dried at room temperature. The concentration of the analyte was ~5–25 pmol/μl. Measurements were performed on a Bruker REFLEX™ MALDI/TOF mass spectrometer using a N_2 laser (337 nm) with a 3-ns pulse width and 107–108 watt/cm2 irradiance at the surface (0.2 mm^2 spot). Spectra were recorded at an acceleration voltage of 28.5 kV in the linear mode, using the delayed extraction facility.

RESULTS

Phytepsin in Grains and Roots—Affinity-purified phytepsin preparation from barley grains typically contains two enzyme forms of approximately 32 + 16 kDa and 29 + 11 kDa and occasionally some higher molecular mass precursors. In earlier studies (9, 10) molecular weight estimations for different chains were based on SDS-PAGE analyses (Fig. 1) and on calculations from the cDNA-derived protein sequence when the N-terminal sequence was known. However, details about the processing of the different polypeptides at the C terminus were not known. To more accurately describe the sizes of the polypeptides resulting from the processing events, we used

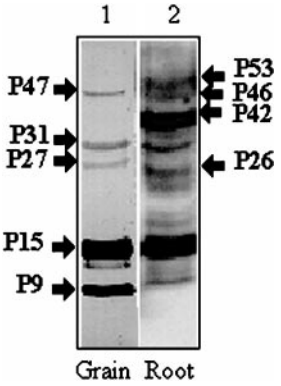

FIG. 1. **Western blot analysis of phytepsin.** *Lane 1*, phytepsin purified from grains, 4 μg of protein; *lane 2*, barley root extract, 48 μg of protein. Samples 1 and 2 were run in the same gel, but sample 1 was developed with 4-chloro-1-naphthol, and sample 2 was developed by the ECL method. Polypeptides from purified phytepsin, whose molecular weights were detected by MALDI-TOF-MS, are indicated on the left, and those from the roots are indicated on the right.

MALDI-TOF-MS for the analysis of reduced processed products that corresponded to the polypeptides constituting the two isoforms of phytepsin from barley grains. The sizes determined for two or three independent phytepsin preparations were 9.2–9.5, 15.3–15.8, 26.4–26.7, 30.5–30.8, and 46–47 kDa. Accordingly, the polypeptides were named P9, P15, P27, P31, and P47 (indicated on the left side of Fig. 1). When we analyzed barley root extracts by Western blotting with the same anti-phytepsin antibody, we detected polypeptides P9, P15, and P31. We also detected a polypeptide of ~26 kDa in roots instead of 27 kDa as previously observed in grains (14), and additional polypeptides of approximately 42, 46, and 53 kDa (P42, P46, and P53, respectively) (Fig. 1, *lane 2*). The similarity among the molecular weights of polypeptides below 31,000 indicates that the processing of phytepsin follows a similar pathway in both grains and roots. P53 and P46 correspond to glycosylated prophytepsin and one-chain phytepsin, respectively. P42 probably corresponds to an additional independent isoenzyme characteristic of the root, as we discuss later in the text. Thus, the roots constitute a good model for studying the expression and processing of phytepsin during the intracellular transport to its final cellular location, the vacuole.

Processing Rate of Phytepsin in Root Cells—The expression and processing of phytepsin were followed by pulse-chase labeling of the roots with [^{35}S]methionine/cysteine. After labeling, the roots were homogenized under nondenaturing conditions, and the processing products were immunoprecipitated with anti-phytepsin antiserum and analyzed by SDS-PAGE (Fig. 2). The precursor P53 appeared during the first 30 min of the pulse, and its half-life was estimated to be 3 h because it was no longer detected 6 h after the beginning of the chase (Fig. 2, *A* and *B*). The primary processing products, P31 and P15, appeared 45 min after the pulse (Fig. 2*A*), and the further processed polypeptides, P26 and P9, were only observed after 24 h of chase (Fig. 2*B*). The processing products of phytepsin were still detected 3 days after the chase (data not shown), indicating a slow turnover rate. The pulse-chase analysis also showed the appearance of an additional protein, P42, 30 min after the pulse. The intensity of this protein remained constant for the entire 24-h chase period, which suggests that it is not related to the proteolytic processing described above.

Brefeldin A Causes Accumulation of the 53-kDa Precursor—To study whether the processing of P53 occurred before prophytepsin reached the Golgi complex, we treated the roots with the fungal antibiotic brefeldin A, which is known to inhibit Golgi-mediated vesicular traffic by disrupting the

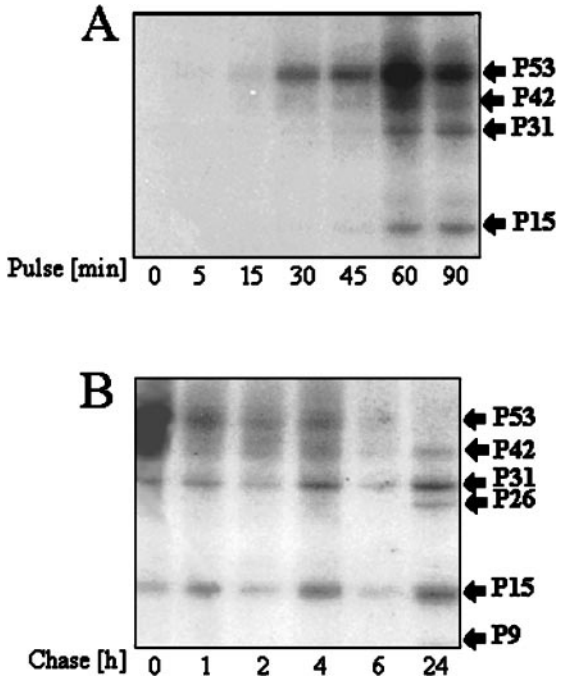

FIG. 2. **Processing of phytepsin in barley root tips.** Roots were labeled with 100 μCi of Pro-mix™ (L-[^{35}S]methionine/cysteine) for the indicated pulse periods, and the phytepsin was immunoprecipitated with anti-phytepsin antibody. Approximately 40 μg of total protein/*lane* was separated by electrophoresis in a gradient (10–20%) SDS-polyacrylamide gel. The gel was dried and fluorographed. In *B*, the pulse time was 2 h.

Golgi apparatus (22). Incubation of root cells with brefeldin A before metabolic labeling followed by pulse-chase experiments showed the accumulation of P53, whereas for the nontreated cells at the same chase times, P53 was processed to the two-chain form P31 + P15 (Fig. 3). Brefeldin A affected the processing for several hours after treatment, and only partial processing of P53 was seen at the 2- and 5-h time points. The experiment with brefeldin A clearly shows that processing of P53, which leads to the formation of the two-chain form of phytepsin, occurs only after the precursor has reached the Golgi complex or has migrated beyond it. Brefeldin A did not have any effect on the P42 polypeptide, further corroborating our assumption that it is independent from the described proteolytic pathway.

P53 Acquires Endo H Resistance—Endo H removes oligomannose but not complex-type *N*-linked glycans from glycoproteins. Processing of oligomannose to complex-type glycans occurs in the Golgi complex, and therefore, resistance to Endo H indicates localization of a glycoprotein at or beyond the Golgi complex. The prophytepsin sequence contains a single *N*-glycosylation site located in P15, and the attached glycans in the mature P31 + P15 form are known to be of the plant complex type (23). During the time course of a 30-min chase, the 53-kDa precursor was partially sensitive to Endo H (detected as a wider band), with a shift of about 2 kDa (Fig. 4). This result shows that the glycans linked to the P53 proform are of the oligomannose type for about 30 min after the beginning of the chase, which corresponds to the time taken for the enzyme to reach the Golgi complex, where the glycans are modified. In contrast, the P15 chain was not sensitive to Endo H digestion at any time point (Fig. 4*B*), indicating that P15 contains only complex-type glycans. Therefore, we conclude that P15 is produced only when prophytepsin has passed the Golgi complex and, most likely, in transit to or within the vacuole.

N-glycans Delay the Proteolytic Processing of Phytepsin—

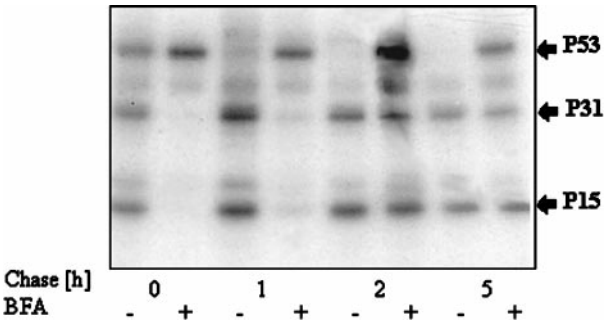

FIG. 3. **Inhibition of phytepsin processing in root tips by brefeldin A.** Roots were preincubated for 1 h with 7 μM brefeldin (*BFA*). This was followed by radiolabeling with 100 μCi of Pro-mix™ (L-[^{35}S]methionine/cysteine) per sample and incubation for the indicated pulse periods. Phytepsin was immunoprecipitated with anti-phytepsin antibody. Approximately 40 μg of total protein/*lane* was separated by electrophoresis in a gradient (10–20%) SDS-polyacrylamide gel. The gel was dried and fluorographed.

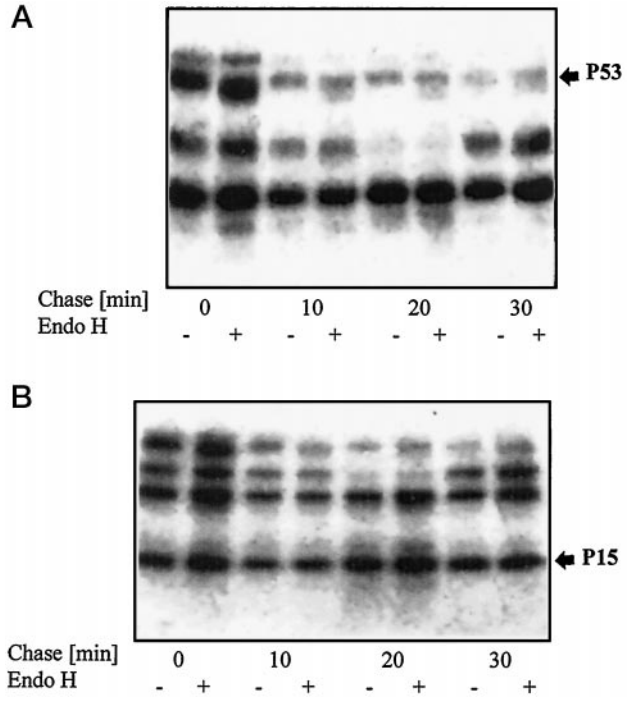

FIG. 4. **Endo H sensitivity of phytepsin from root tips after pulse-chase analysis.** Roots were treated as described in Fig. 2. After immunoprecipitation with anti-phytepsin antibody, half of the sample was digested with Endo H, and the other half was mock-digested. Approximately 40 μg protein/*lane* was separated by electrophoresis in a gradient (*A*) 10–20% or (*B*) 10% SDS-polyacrylamide gel. The gels were dried and fluorographed.

Plant APs contain a conserved utilized glycosylation site in their plant-specific insert (10). Since it has been suggested that the plant-specific insert might be important for transport to the vacuole (13), we investigated the importance of glycans for the intracellular transport. We found that incubation of root cells with the *N*-glycosylation inhibitor tunicamycin did not inhibit processing of P53 (Fig. 5), but that the processing of P53 to P31 + P15 and P26 + P9 was actually accelerated when glycosylation was inhibited. These results suggest that the glycan moiety of phytepsin protects the enzyme from premature proteolytic cleavage in the Golgi apparatus.

Expression and Purification of Phytepsin Produced in Sf9 Cells—To enable a closer study on the processing pattern of phytepsin, we developed a recombinant expression method for this enzyme. Although several APs, including pepsinogen (24),

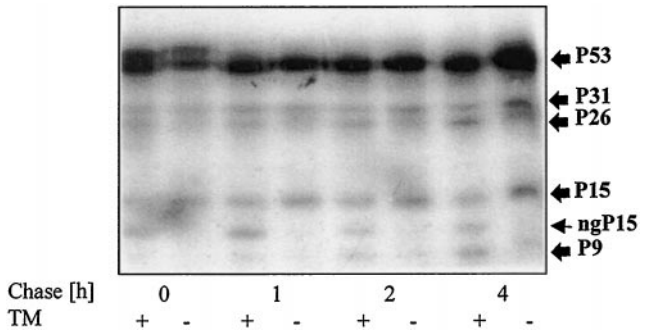

FIG. 5. **Processing of phytepsin in root tips in the presence of tunicamycin A_1.** Roots were preincubated for 1 h with 10 μg/ml tunicamycin A_1 homologue (*TM*), radiolabeled with 100 μCi of Pro-mix™ (L-[^{35}S]methionine/cysteine)/sample. The phytepsin was immunoprecipitated with anti-phytepsin antibody. *ngP15*, nonglycosylated P15. Approximately 40 μg of total protein/*lane* was separated by electrophoresis in a gradient (10–20%) SDS-polyacrylamide gel. The gel was dried and fluorographed.

procathepsin D (5, 6) and phytepsin,[2] have been produced in bacterial expression systems, a general problem with these methods has been the very low yield of correctly folded product. Therefore, we chose a baculovirus-infected insect cell expression method for this study. Sf9 cells were infected by a recombinant baculovirus genome containing the complete coding region of preprophytepsin. After incubation for 4 days, a prominent polypeptide of 53 kDa was found in the medium, as analyzed by Western blotting using anti-phytepsin antiserum (not shown). After pepstatin-agarose column chromatography followed by ion exchange chromatography, a typical yield of about 0.5 mg of purified protein was obtained from 1 liter of the cell medium. MALDI-TOF-MS analysis revealed two proteins of 52,847 and 53,062 Da (Fig. 6), which migrated in SDS-PAGE as one broad band at 53 kDa (Fig. 7*A*, *lane 1*). N-terminal sequencing gave one unambiguous sequence of EAEGLVRIAL (Fig. 8). This N-terminal sequence is similar to that previously obtained from *in vitro* expression in the presence of canine pancreatic microsomes (14), which indicates that Sf9 cells are able to cleave the signal sequence from this plant protein to produce a secreted recombinant prophytepsin (rP53). In an isoelectric focusing gel, rP53 migrated to a pI of ~5.3, which was identical to that observed for phytepsin purified from barley grains (data not shown).

The molecular weight for rP53 predicted from the cDNA is 51,779. The higher molecular weight observed by MALDI-TOF-MS is due to the presence of oligosaccharides, confirmed by positive staining observed using the periodic acid Schiff method to detect the protein in the gel and indicates that the single potential glycosylation site is occupied. The observed main peak at 52,847 is consistent with the presence on prophytepsin of an *N*-linked oligosaccharide with the structure Man3GlcNAcFucGlcNAc characteristic of proteins expressed in insect cells (predicted molecular weight of 52, 839 (25), assuming the presence of one sodium atom). The additional signals at 53,062 and 53,273 are probably due to the addition of one or two matrix molecules (sinapinic acid) with the concomitant loss of water (M + nx206) or to the presence of larger oligosaccharide chains.

Autoproteolytic Processing of Recombinant Phytepsin in Vitro—To study the capability of prophytepsin for autoproteolytic processing, we incubated rP53 in buffers over the pH range 3.7–6.5 at 37 °C and removed the samples at various time points. At pH 3.7, the processed polypeptides of 36 and 17 kDa (rP36 and rP17, respectively) were detected after 7 min of

[2] J. Kervinen, unpublished information.

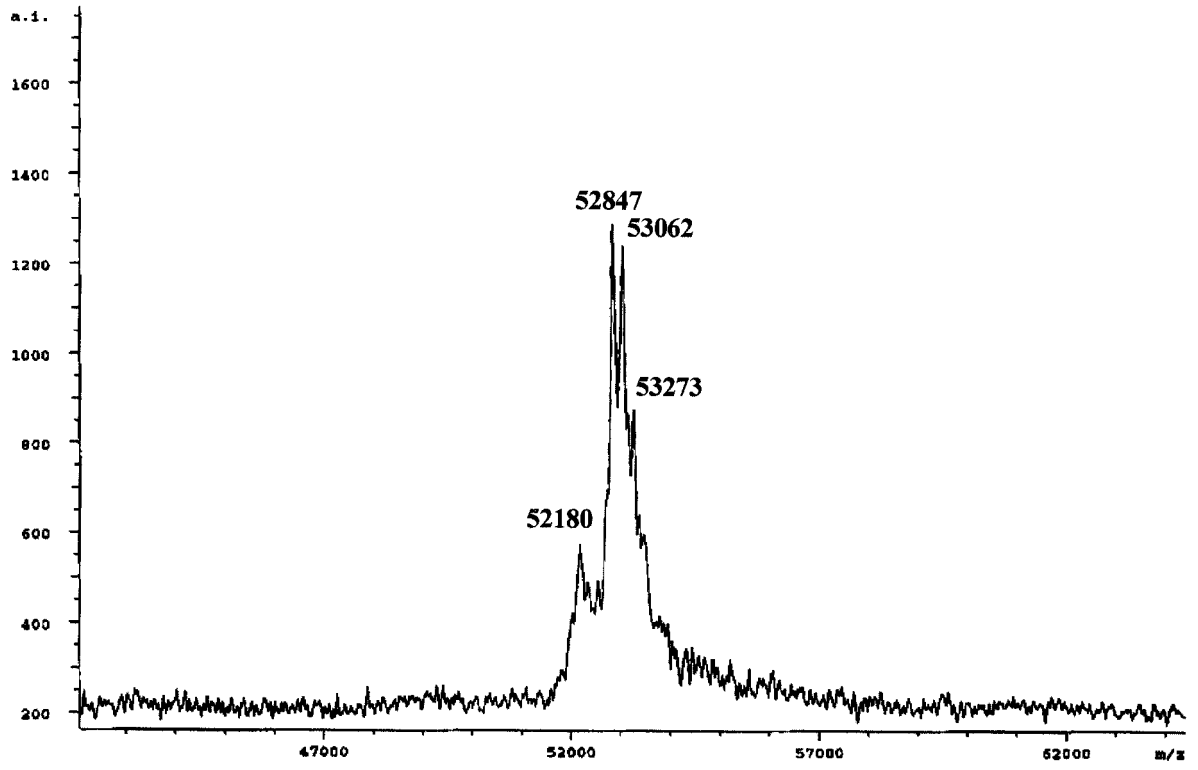

FIG. 6. **MALDI-TOF-MS analysis of the precursor rP53.** *a.i.*, relative intensity.

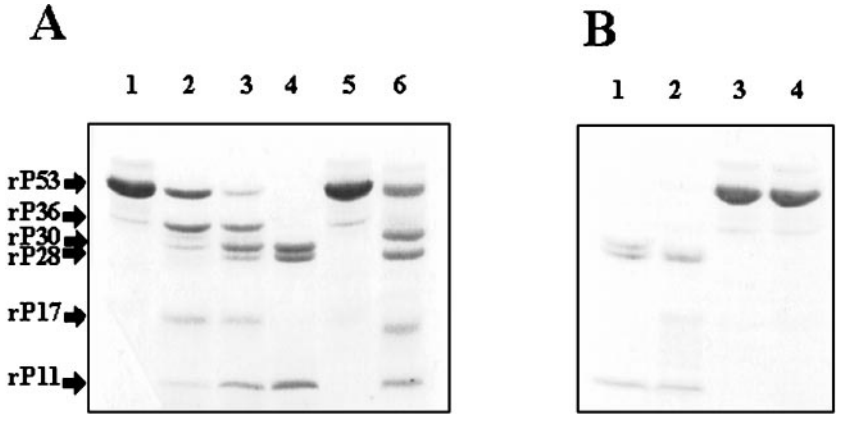

FIG. 7 **Autoproteolytic processing of recombinant prophytepsin *in vitro*.** *A*, purified prophytepsin from insect cell medium (*lane 1*) was mixed with an equal volume of 0.2 M sodium lactate, pH 3.7, and incubated at 37 °C. Samples were removed after 7, 15, and 60 min (*lanes 2–4*, respectively); *lane 5*, incubation for 60 min with 50 μM pepstatin; *lane 6*, phytepsin purified from barley grains (3 μg). *B*, incubation of prophytepsin for 90 min at pH 3.7, 4.5, 5.5, and 6.5 (*lanes 1–4*, respectively). All samples were analyzed by electrophoresis in 20% SDS-polyacrylamide gels and stained with Coomassie Brilliant Blue.

incubation, indicating rapid cleavage of prophytepsin to a two-chain form (rP36 + rP17) (Fig. 7*A*, *lane 2*). The two polypeptides were further processed stepwise to the final products of 28 and 11 kDa (rP28 and rP11, respectively). At pH 3.7, processing of rP30 to rP28 occurred slower than at pH 4.5, whereas processing of rP17 to rP11 was much slower at pH 4.5 than at pH 3.7. Processing was severely inhibited at pH values above 4.5 (Fig. 7*B*). The addition of the AP inhibitor pepstatin to the incubation buffer prevented all processing and showed that the processing was of an autoproteolytic nature (Fig. 7*A*, *lane 5*). Polypeptides rP28 and rP11 were the final processing products over the pH range 3.7–4.5 and up to 90 min of incubation; furthermore, in some extended incubations at pH 4.5, no additional processing was observed after 5 h. The sizes of the proteolytic processing products were very similar to those of the polypeptides present in phytepsin purified from barley grains (Fig. 7*A*, *lane 6*). To obtain detailed information on the cleavage sites, we separated the processing products of rP53 by SDS-PAGE, electroblotted them onto a PVDF membrane, and subjected several of them to N-terminal sequencing. rP36, rP30, and rP28 each gave one sequence, whereas rP17 and rP11 both gave two or three almost identical sequences, indicating slight heterogeneity in these processing sites. The sequences and processing sites are illustrated in Fig. 8. The molecular weights of the polypeptides resulting from the autolysis of rP53 were confirmed by MALDI-TOF-MS.

Specific Activity of Recombinant Phytepsin—To test the proteolytic efficiency of the purified and autoproteolytically processed recombinant phytepsin, we subjected prophytepsin to autocatalytic activation at pH 4.0 for 1 h at 37 °C and assayed the proteolytic activity against hemoglobin at 37 °C. The observed mean value for purified recombinant phytepsin from three individual expressions was 628 units/mg (range 456–788 units/mg), whereas the purified enzyme from grains gave a

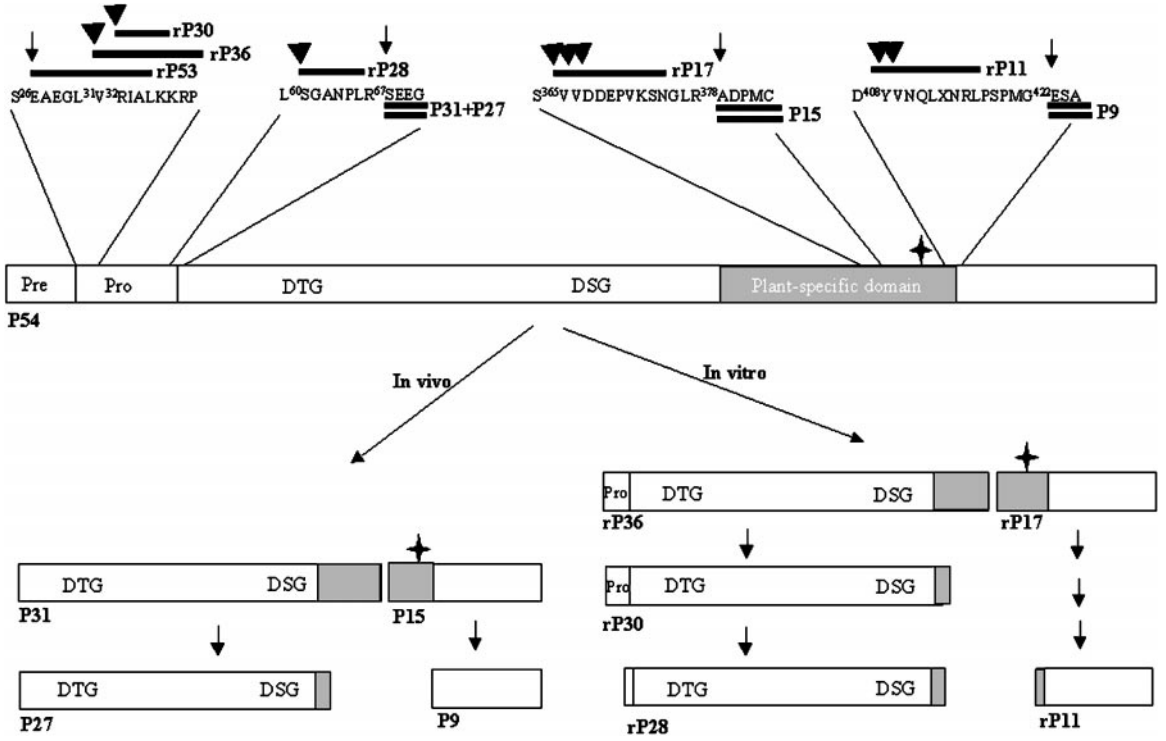

FIG. 8. **Proteolytic processing scheme of phytepsin.** ↓, *in vivo* cleavage site; ▼, *in vitro* autocatalytic cleavage site. *Bold Double underlining* denotes the N-terminal polypeptides from *in vivo* processing (10) and overlining N-terminal polypeptides from *in vitro* autoprocessing. *Bold cross*, glycosylation site. Approximate locations of catalytic residues in the active site are marked by *DTG*/*DSG*. The plant-specific domain is *shaded*.

value of 882 units/mg. These values are consistent with the previously reported value for affinity-purified phytepsin from grains (range 534–668 units/mg at 30 °C) (9) and indicate that expression of recombinant phytepsin in baculovirus-infected insect cells yields a correctly folded and active enzyme.

DISCUSSION

Phytepsin is synthesized and translocated into the rough endoplasmic reticulum as a preproenzyme of 54 kDa, according to the cDNA sequence, where it becomes *N*-glycosylated at its single glycosylation site. The enzyme undergoes several proteolytic cleavages to produce the mature two-chain forms present in barley grains, roots, and other tissues:

1) The first processing step consists of the removal of the signal sequence of 25 amino acid residues upon entering the endoplasmic reticulum, yielding a product of 51,779 Da. Even though this cleavage has not been studied *in vivo* in barley, the cleavage site of the signal sequence occurs between Ser-25 and Glu-26, as has been detected *in vitro* using microsomal membranes (14) or as shown here for the recombinant preprophytepsin from baculovirus-infected insect cells (Fig. 8). Concomitantly, the *N*-glycan $Glc_3Man_9GlcNAc_2$ is transferred onto the single glycosylation site of the enzyme and is further processed to the structure Manα3,6(±Manα3,6)(Xylβ2)Manβ4 GlcNAcβ4(Fucα3)GlcNAc in barley (23), resulting in an increase in molecular weight to ~53,000. Alternatively, when the enzyme is expressed in Sf9 insect cells, the oligosaccharide is processed to the structure Manα6(Manα3)Manβ4GlcNAcβ4 (Fucα6)GlcNAc, which corresponds to the glycosylated precursor rP53 (Fig. 6).

2) The following processing steps include the removal of the N-terminal propeptide sequence of 41 residues and the formation of a two-chain form of the enzyme. The two polypeptides (P31 + P15) were detected 1 h after synthesis. Considering the N-terminal sequence of P15 (Fig. 8), a molecular weight of 14,046 is predicted from the cDNA sequence, and the addition of the glycan at the single glycosylation site accounts for the observed extra 1.2 kDa. Thus, no processing at the C terminus of this subunit occurred. However, if a single proteolytic cleavage had occurred on the 53-kDa polypeptide to produce P15, the larger subunit should be 38 kDa, taking into account the N terminus of P31 (Fig. 8). This difference in molecular weight can be attributed to proteolytic cleavage of a 5-kDa peptide upstream of Ala-378, in the middle of the coding sequence of the enzyme. This processing step is inhibited by brefeldin A and P53 accumulates. Brefeldin A is known to inhibit transport beyond the Golgi complex, so the results suggest that the proteolytic processing occurs after phytepsin has passed the Golgi, most probably in the vacuole. Further corroborating these results, P53 is sensitive to Endo H for 30 min, the time required for P53 to reach the Golgi complex and for the glycans to be processed from oligomannose to complex type by Golgi glycosidases and glycosyltransferases, whereas P15 is always Endo H-resistant, indicating that it is produced only after the glycans have been processed, and thus, the phytepsin has passed the Golgi complex.

3) Finally, P31 and P15 are further processed, resulting in polypeptides P27 and P9, respectively. P27 results from P31 after C-terminal processing, since the N termini of P31 and P27 are identical (Fig. 8). P15 is probably processed only from the N-terminal side, since removal of 44 residues (~ 5 kDa) including the *N*-linked glycosylation site with attached glycans decreases the size to P9. This processing step occurs only 24 h after synthesis *in vivo*.

In addition to the polypeptides described in the processing pattern described above, a polypeptide of approximately 42 kDa is clearly visible in a Western blot from roots (Fig. 1) and in the pulse-chase experiments (Figs. 2–5). However, since the intensity of P42 remains constant during the pulse-chase experiments, we propose that P42 is not part of the processing scheme of P53, but rather, that it probably constitutes a different AP-like isoenzyme from the root. Further supporting a hypothesis that plant tissues contain several AP-like enzymes,

Nakano *et al.* (26) recently showed that chloroplast nucleoids from tobacco cells contain a DNA-binding protein of 41 kDa with clear sequence homology to the AP family, including the conserved active-site residues Asp-Thr-Gly/Asp-Ser-Gly. Chen and Foolad (27) also reported the cloning and characterization of an AP-like proteinase of ~45 kDa (nucellin) from barley ovaries, which is abundantly expressed after pollination. Furthermore, a putative AP is encoded by part of the *BARE-1* retroelement in the barley genome (28). Interestingly, the aforementioned AP-like enzymes do not contain a plant-specific domain that in phytepsin is cleaved and results in the two-chain enzyme form (Fig. 8).

Plant APs contain a conserved occupied *N*-glycosylation site within the plant-specific insert and we have investigated the importance of the glycosylation for the intracellular transport and processing of the enzyme. The results obtained with tunicamycin suggest that the glycans are not essential for the transport or processing of phytepsin, since proteolytic processing occurred despite the presence of this compound. In fact, the rate of processing was accelerated by treatment with tunicamycin, suggesting that the glycans may have a role in protecting the enzyme from premature proteolytic cleavage.

The recombinant prophytepsin (rP53) expressed in baculovirus-infected insect cells was detected in the medium as a glycosylated proenzyme. The enzyme contained glycans of the insect complex type, indicating that rP53 had passed the Golgi complex before secretion outside the cell. It is known that cathepsin D and many other animal proteins are targeted largely to the lysosome via the mannose-6-phosphate receptor pathway (3). Such a targeting mechanism to the plant vacuoles has not been described. Apparently, phytepsin was secreted from the insect cells because it did not contain the appropriate intracellular targeting signal functional in insect cells. It is also possible that phytepsin was secreted because the cells could not hold the large amount of foreign protein they were producing.

When recombinant prophytepsin was incubated at pH 3.7–4.5, it underwent autoprocessing, even though the cleavage sites were distinct from those occurring *in vivo* (for a comparison, see Fig. 8). Results from N-terminal sequencing and the decrease in molecular masses revealed that 1) the 53-kDa glycosylated prophytepsin was autoprocessed *in vitro* to rP36 and rP17, with cleavage occurring around Val-365; 2) rP36 was further processed from both N- and C-terminal sides to yield the final proteolytic product, rP28; and 3) the glycosylated polypeptide rP17 was processed only from its N-terminal side to rP11 (Fig. 8).

The recombinant phytepsin possesses features similar to those of phytepsin purified from barley grains (9, 16). It hydrolyzes peptide bonds that usually contain at least one hydrophobic residue in either side of the bond to be cleaved (Fig. 8), and the optimal pH for the hydrolytic activity is 3.7–4.5 (Fig. 7). Furthermore, the enzymatic efficiency of the recombinant phytepsin on hemoglobin was almost as high as that measured for phytepsin purified from grains. Additionally, in an ongoing related study, we have crystallized rP53, and a preliminary x-ray analysis extending to 2.4 Å resolution shows that the overall fold of prophytepsin, excluding the plant-specific domain, is similar to that of mammalian APs.[3]

Although the sizes of the autocatalytic processing intermediates and final products of recombinant phytepsin closely resemble the *in vivo* P31 + P15 and P27 + P9 forms of barley grain phytepsin, the cleavage products formed *in vitro* are not exactly the same. rP28 still retains seven residues belonging to the prosequence, and rP11 contains 13–14 extra residues in the N-terminal side. In addition, the very hydrophilic and, thus, unfavorable residues, Arg-Ser in the N terminus of P31 and P27, Arg-377 in P15, and Gly-Glu in P9 (Fig. 8), are flanking the final maturation sites in phytepsin purified from barley grains. Therefore, although *in vitro* processing of rP53 results in an active enzyme, the completion of maturation *in vivo* probably requires other proteinase/exopeptidase(s).

Both autocatalytic and heterocatalytic processing and activation mechanisms are known for APs. For example, the activation of mammalian lysosomal procathepsin D, the closest counterpart to phytepsin, involves cysteine proteases in the lysosomes (4). However, *in vitro* studies show that procathepsin D undergoes a pH-dependent intramolecular proteolysis that removes 26 residues from the 44-residue propart, yielding an active one-chain enzyme, pseudocathepsin D (5, 6). On the other hand, when a mutant of procathepsin D that was unable to autoactivate itself to pseudocathepsin D *in vitro* was expressed in mouse cells, it nevertheless was transported to the lysosomes and was processed normally to the mature two-chain enzyme (7). Based on several processing studies on intracellular APs, it thus seems likely that alternative activation mechanisms, including autocatalytic and heterocatalytic steps, exist for intracellular APs, as we have shown for phytepsin in the present study. These mechanisms depend on the pH and processing endo- and exopeptidases present in the particular intracellular compartments traversed by the AP along its route to its final location in the cell.

Acknowledgments—We gratefully acknowledge the scientific advice of Dr. Pedro Fevereiro and Lucinda Neves and Anne Arthur for her editorial assistance.

[3] J. Kervinen and A. Zdanov, unpublished information.

REFERENCES

1. Davies, D. (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19,** 189–215
2. Rawlings, N. D., and Barrett, A. L. (1995) *Methods Enzymol.* **248,** 105–120
3. Metcalf, P., and Fusek, M. (1993) *EMBO J.* **12,** 1293–1302
4. Samarel, A. M., Ferguson, A. G., Decker, R. S., and Lesch, M. (1989) *Am. J. Physiol.* **257,** C1069–C1079
5. Conner, G. E., and Richo, G. (1992) *Biochemistry* **31,** 1142–1147
6. Beyer, B. M., and Dunn, B. M. (1996) *J. Biol. Chem.* **271,** 15590–15596
7. Richo, G. R., and Conner, G. E. (1994) *J. Biol. Chem.* **269,** 14806–14812
8. Barrett, A. J. (1997) *Eur. J. Biochem.* **250,** 1–6
9. Sarkkinen, P., Kalkkinen, N., Tilgmann, C., Siuro, J., Kervinen, J., and Mikola, L. (1992) *Planta* **186,** 317–323
10. Runeberg-Roos, P., Törmakangas, K., and Östman, A. (1991) *Eur. J. Biochem.* **202,** 1021–1027
11. Cordeiro, M. C., Xue, Z. T., Pietrzak, M., Pais, M. S., and Brodelius, P. E. (1994) *Plant Mol. Biol.* **24,** 733–741
12. Asakura, T., Watanabe, H., Abe, K., and Arai, S. (1995) *Eur. J. Biochem.* **232,** 77–83
13. Guruprasad, K., Törmakangas, K., Kervinen, J., and Blundell, T. L. (1994) *FEBS Lett.* **352,** 131–136
14. Runeberg-Roos, P., Kervinen, J., Kovaleva, V., Raikhel, N. V., and Gal, S. (1994) *Plant Physiol.* **105,** 321–329
15. Marttila, S., Jones, B. L., and Mikkonen, A. (1995) *Physiol. Plant.* **93,** 317–327
16. Kervinen, J. (1998) in *The Handbook of Proteolytic Enzymes* (Barrett, A. J., Woessner, F. F., and Rawlings, N., eds) Academic Press, Inc. London, in press
17. Runeberg-Roos, P., and Saarma, M. (1998) *Plant J.* **15,** 139–145
18. Pinkas-Kramarski, R., Shelly, M., Glathe, S., Ratzkin, B., and Yarden, Y. (1996) *J. Biol. Chem.* **271,** 19029–19032
19. Laemmli, U. K. (1970) *Nature* **227,** 680–685
20. Costa, J., Grabenhorst, E., Nimtz, M., and Conradt, H. S. (1997) *J. Biol. Chem.* **272,** 11613–11621
21. O'Reilly, D. R., Miller, L. K., and Luckow, V. A. (1994) *Baculovirus Expression Vectors: A Laboratory Manual*, Oxford University Press, New York
22. Satiat-Jeunemaitre, B., Cole, L., Bourett, T., Howard, R., and Hawes, C. (1996) *J. Microsc. (Oxf.)* **181,** 162–177
23. Costa, J., Ashford, D. A., Nimtz, M., Bento, I., Frazão, C., Esteves, C. L., Faro, C. J., Kervinen, J., Pires, E., Veríssimo, P., Wlodawer, A., and Carrondo, M. A. (1997) *Eur. J. Biochem.* **243,** 695–700
24. Lin, X., Wong, R. N. S., and Tang, J. (1989) *J. Biol. Chem.* **264,** 4482–4489
25. Grabenhorst, E., Hofer, B., Nimtz, M., Jäger, V., and Conradt, H. S. (1993) *Eur. J. Biochem.* **215,** 189–197
26. Nakano, T., Murakami, S., Tsubasa, S., Yoshida, S., Yamada, Y., and Sato, F. (1997) *Plant Cell* **9,** 1673–1682
27. Chen, F., and Foolad, M. R. (1997) *Plant Mol. Biol.* **35,** 821–831
28. Manninen, I., and Schulman, A. H. (1993) *Plant Mol. Biol.* **22,** 829–846